Прикладная математика (4 семестр)
Лабораторная работа 2
"Методы градиентного спуска"

Александров Даниил, Ершов Александр

ITMO UNIVERSITY

1 Описание

Исследование градиентного спуска

2 Цели

Реализовать методы градиентного спуска с разными алгоритмами задания шага спуска. Реализовать метод сопряженных направлений

3 Задачи

Реализация методов поиска минимума, собрать данные о количестве итераций циклов, проанализировать траекторию поиска, проианализировать сходимость в зафисимости от выбора точки начального приближения, дать оценку эффективности методов

Часть I

Градиентный спуск

Общая идея алгоритма описывается следующим образом:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

В зависимости от того, как задается параметр α_k, существует несколько возможных реализаций метода.

1.1 Метод постоянного шага

В качестве параметра α_k выбираем постоянное число

1.2 Метод дробления шага

В качестве параметра α_k выбираем постоянное число

1.3 Метод золотого сечения при выборе шага

В качестве параметра α_k выбираем такое число, что сечение $\phi(\alpha_k) = f(x_k - \alpha_k \nabla f(x_k)) \rightarrow min$. Причем, поиск такого параметра осуществляется при помощи метода золотого сечения.

1.4 Метод Фибоначчи при выборе шага

В качестве параметра α_k выбираем такое число, что сечение $\phi(\alpha_k) = f(x_k - \alpha_k \nabla f(x_k)) \rightarrow min$. Причем, поиск такого параметра осуществляется при помощи метода Фибоначчи.

В последних двух методах промежуток $[a, b]$, из которого определяется шаг спуска, задается пользователем. Все алгоритмы работают до тех пор, пока не выполнится условие останова:

$$||(x_{k+1} - x_k)|| + |f(x_{k+1}) - f(x_k)| \leq \varepsilon$$

Заметим, что найденная в результате того или иного метода точка лишь гарантирует условие $\nabla f(x_k) = 0$, что говорит о том, что желательно провести дополнительное исследование с целью ее класификации.

Ввиду того, что далеко не каждая функция является выпуклой (тем более строго выпуклой), говорить о сходимости в общем смысле тяжело. Для методов выше определены оценки сходимости, исходя из свойств строгой выпуклости:

$$||(x_{k+1} - x_k)|| \geq \frac{M - m}{M + m}||(x_k - x_*)||$$

, где M и m - оценки наибольшего и наименьшего собственных чисел матрицы $H(x)$

Для методов 1.1 и 1.2 можно потребовать условие Липшица, тогда о сходимости можно будет рассуждать как для методов 1.3 и 1.4.

1.5 Метод Флетчера-Ривса

В качестве параметра α_k выбираем такое число, что сечение $\phi(\alpha_k) = f(x_k - \alpha_k \nabla f(x_k)) \rightarrow min$. Причем, поиск такого параметра осуществляется при помощи метода Фибоначчи.

Приведем некоторые результаты поиска для функций ниже:

Рис. 1: Постоянный

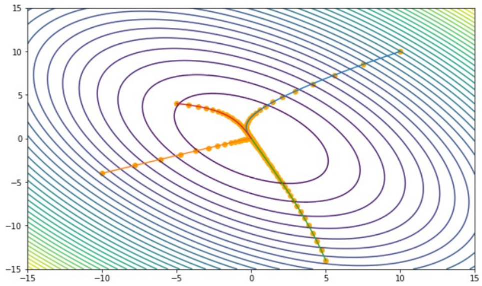


Рис. 2: Дробление

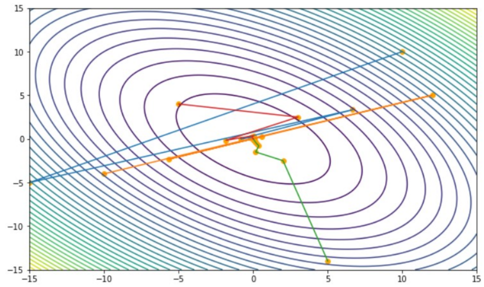


Рис. 3: Золотое сечение

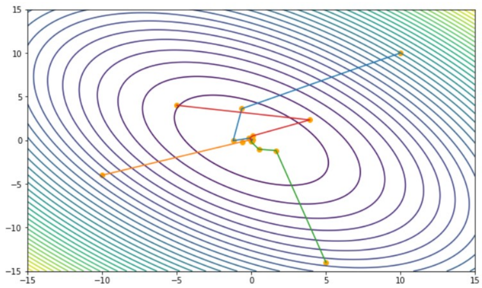
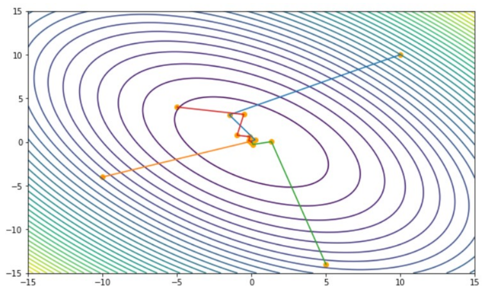


Рис. 4: Фибоначчи



$$f(x, y) = x^2 + xy + y^2$$

Рис. 5: Постоянный

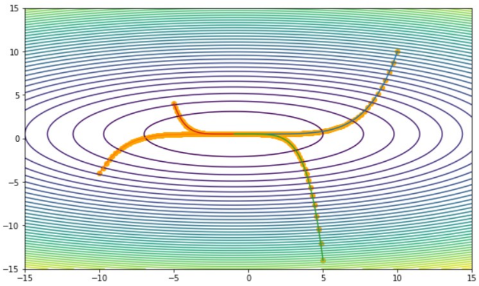


Рис. 6: Дробление

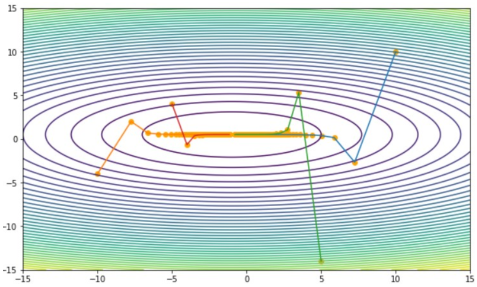


Рис. 7: Золотое сечение

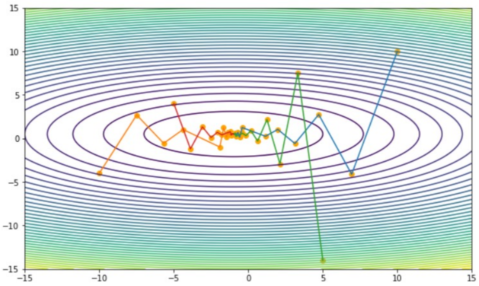
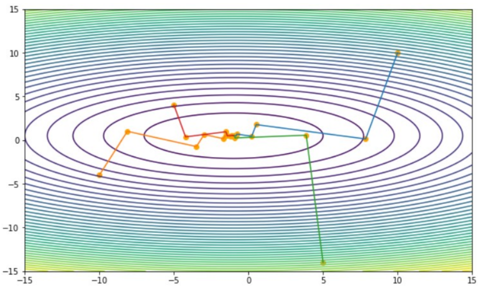


Рис. 8: Фибоначчи



$$g(x,y) = \frac{(x+1)^2}{4} + \frac{(2y-1)^2}{3} + 1$$

Рис. 9: Постоянный

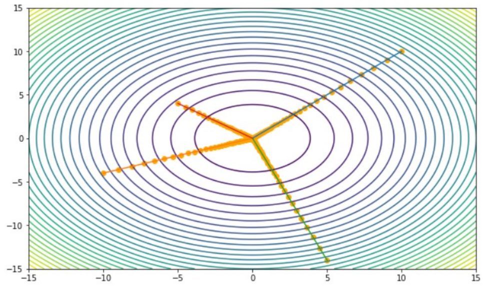


Рис. 10: Дробление

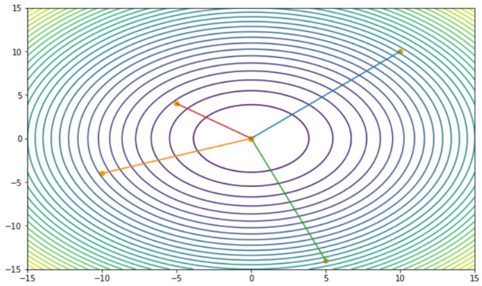


Рис. 11: Золотое сечение

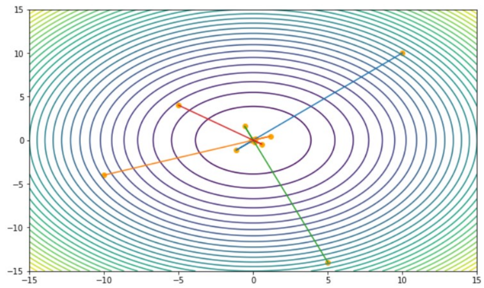
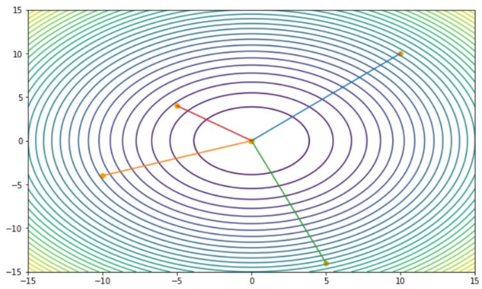


Рис. 12: Фибоначчи



$$h(x, y) = x^2 + y^2 + y + 2$$

Рис. 13: Градиент

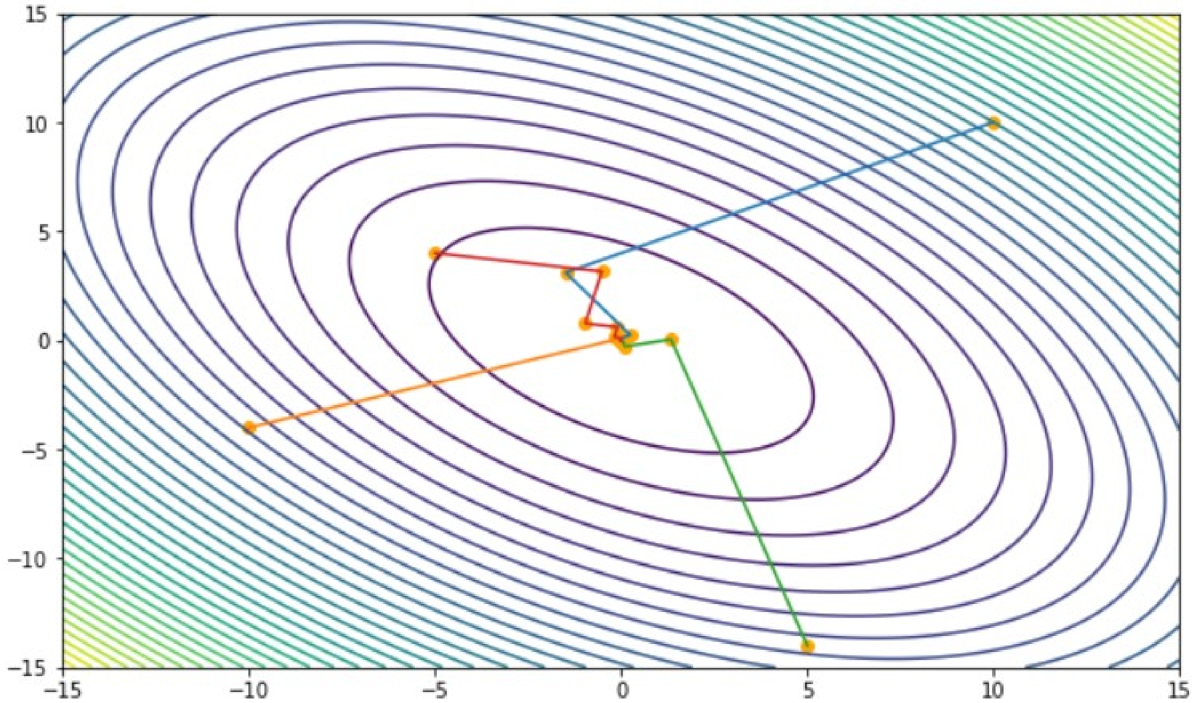


Рис. 14: Флетчер-Ривс

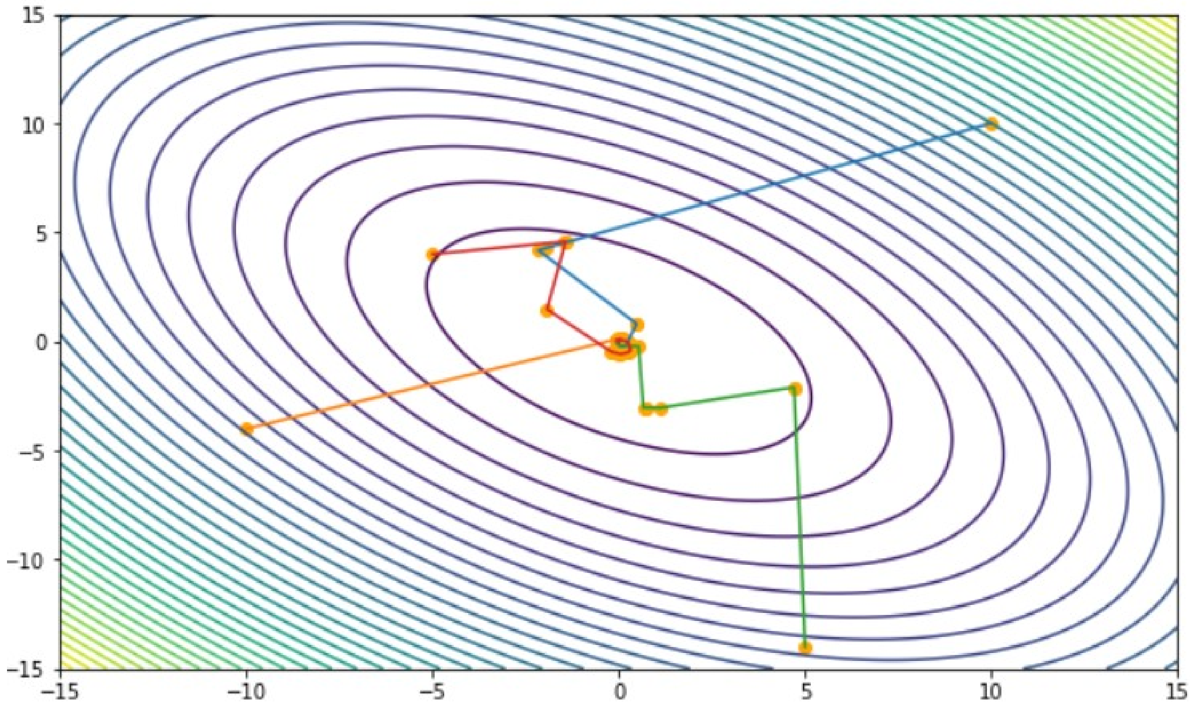


Рис. 15: Градиент

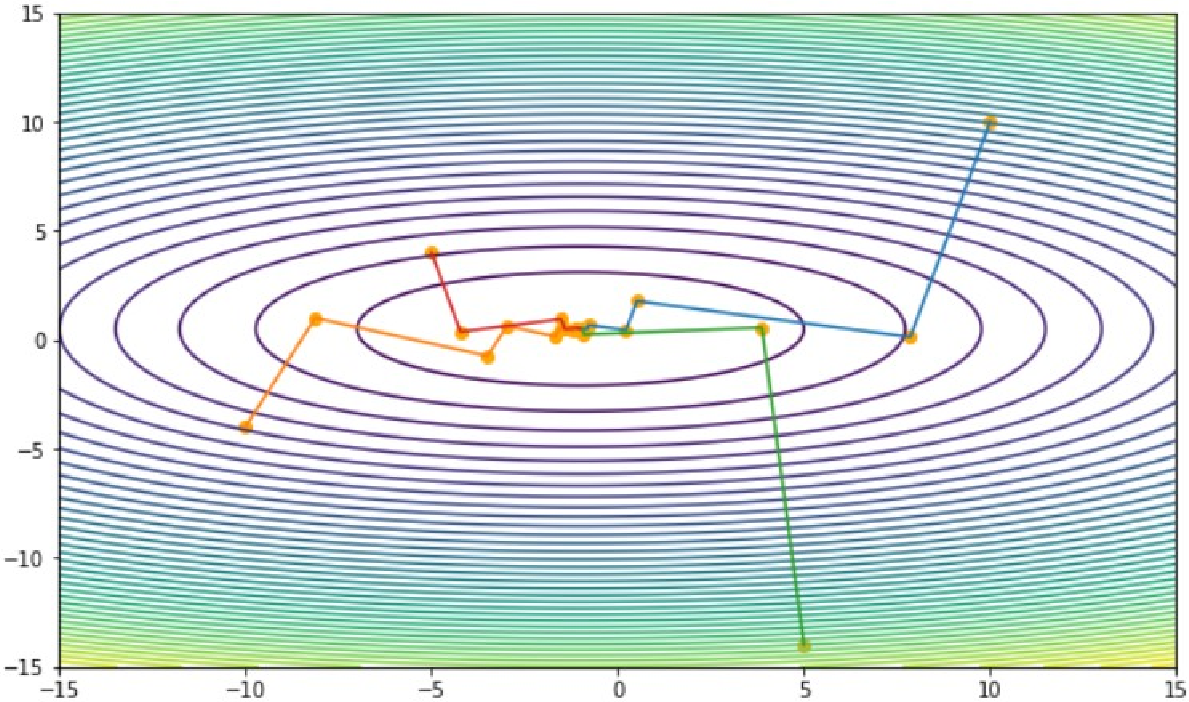


Рис. 16: Флетчер-Ривс

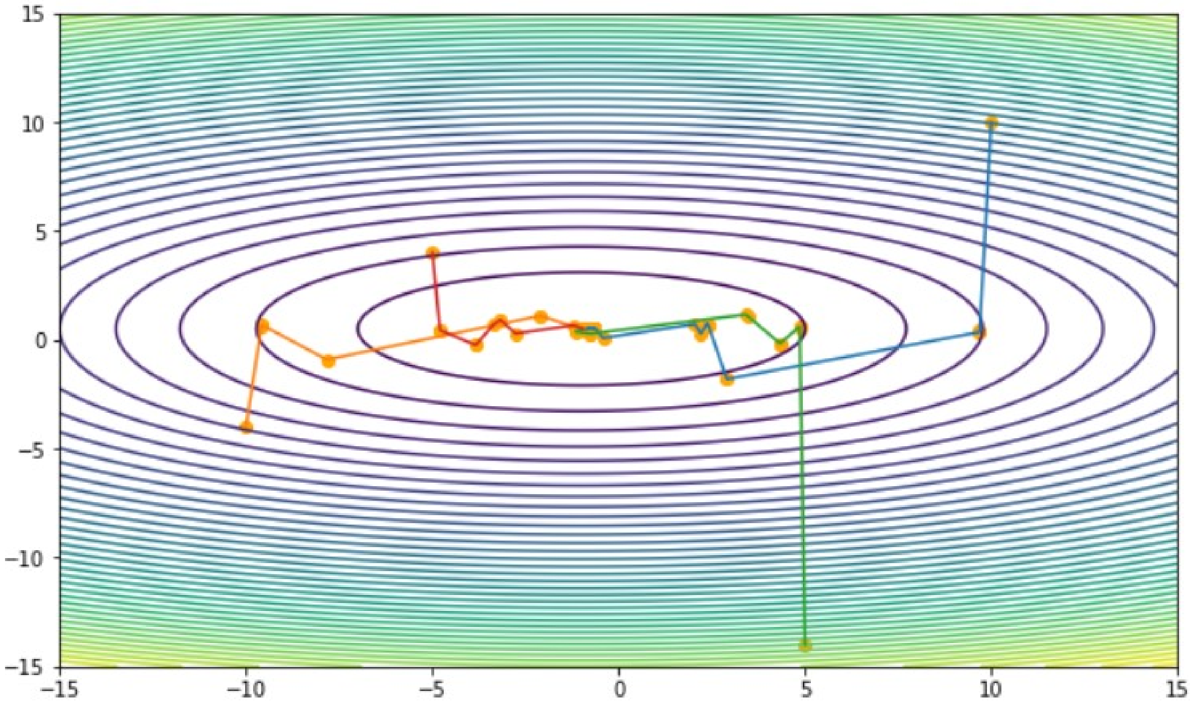


Рис. 17: Градиент

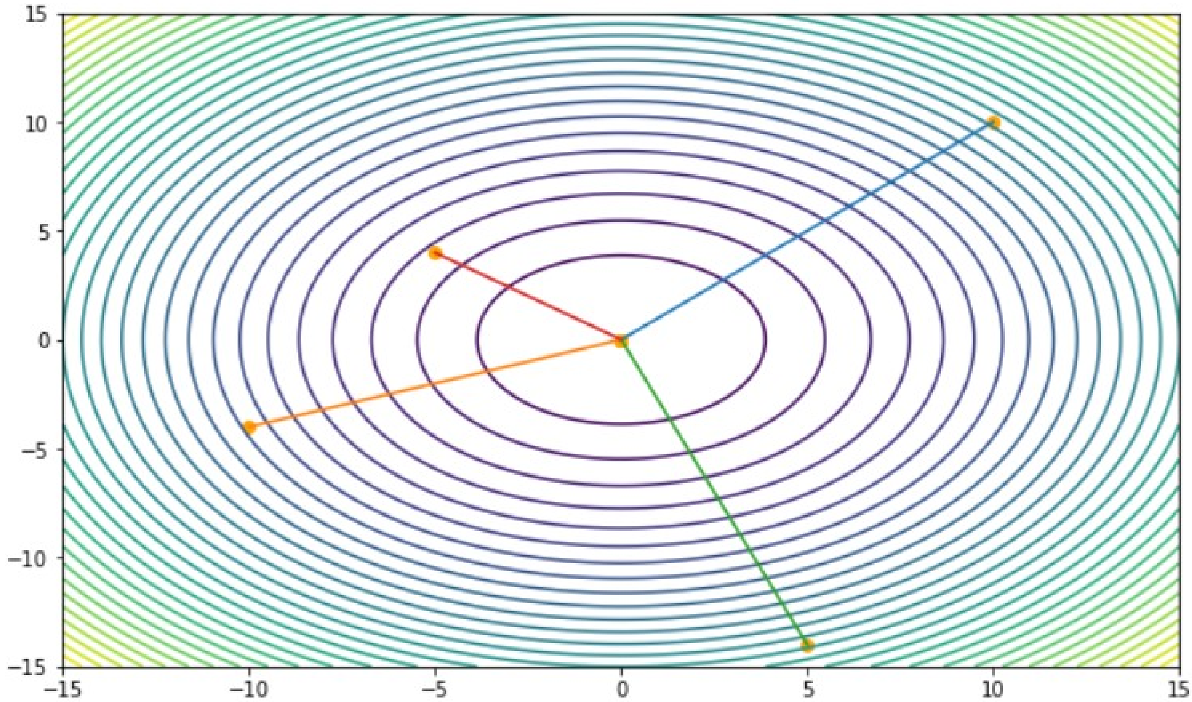
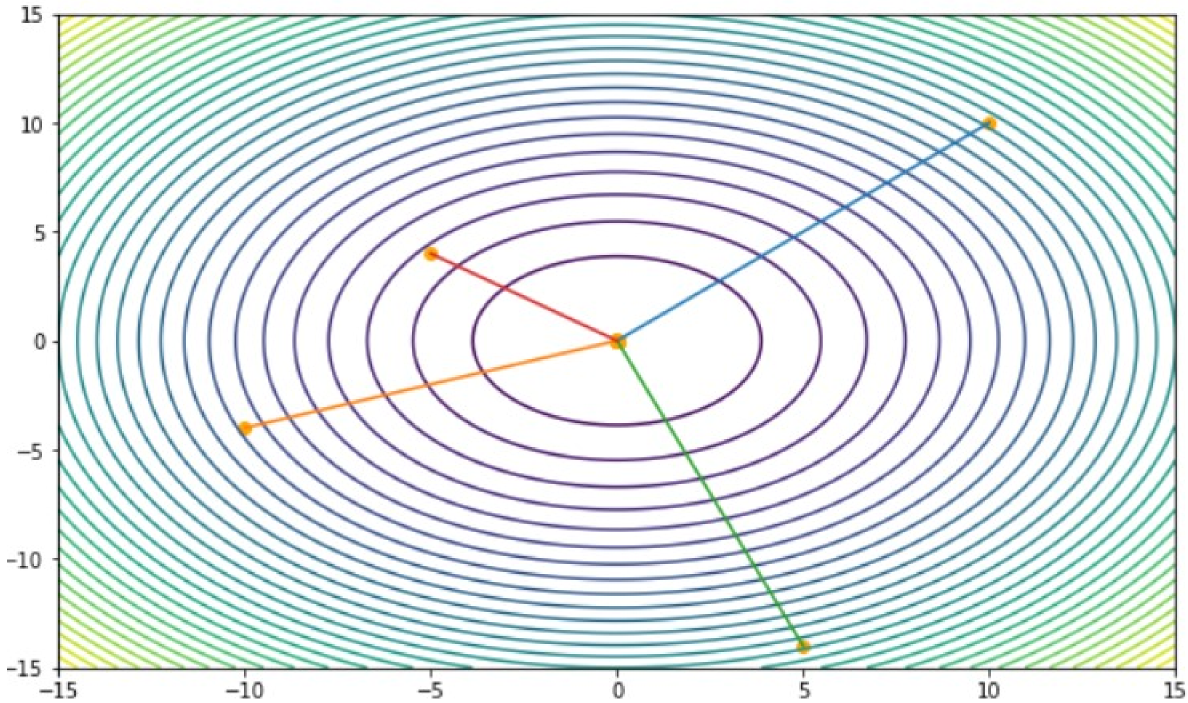


Рис. 18: Флетчер-Ривс



Часть II

Выводы

Градиентный спуск при удачно выбранном постоянном шаге обладает высокой скоростью сходимости. При использовании методов наискорейшего поиска траектория описывает в некотором смасле касательные к линиям уровня, проходящим через точку приближения. (Поговаривают, что если для квадратичных форм в качестве точек начального приближения взять такие, что их координаты совпадают с координатами собственных векторов, то градиентный спуск отработает за один шаг гарантированно).

Методы сопряженного градиента имеют схожую траекторию, что и методы наискорейшего спуска, однако, они более "крутые"в смысле наклона.

Градиентый спуск может не сойтись, если для анализа взять функцию со слишком высокими "стенками"и узким плато. Для того, чтобы обойти эту проблему, можно применить масштабирование.